Relationships among Differential, Truncated Differential, Impossible Differential Cryptanalyses against Word-Oriented Block Ciphers like Rijndael, E2

Makoto Sugita[1], Kazukuni Kobara[2], Kazuhiro Uehara[1], Shuji Kubota[1], Hideki Imai[2]
[1] NTT Wireless Systems Innovation Laboratory, Network Innovation Laboratories,
1-1 Hikari-no-oka, Yokosuka-shi, Kanagawa, 239-0847 Japan
E-mail: {sugita, uehara, kubota}@wslab.ntt.co.jp
[2] Institute of Industrial Sciences, The University of Tokyo
Roppongi, Minato-ku, Tokyo 106-8558, Japan
E-mail:{kobara, imai}@imailab.iis.u-tokyo.ac.jp

Abstract

We propose a new method for evaluating the security of block ciphers against differential cryptanalysis and propose new structures for block ciphers. To this end, we define the word-wise Markov (Feistel) cipher and random output-differential (Feistel) cipher and clarify the relations among the differential, the truncated differential and the impossible differential cryptanalyses of the random output-differential (Feistel) cipher. This random output-differential (Feistel) cipher model uses a not too strong assumption because denying this approximation model is equivalent to denying truncated differential cryptanalysis. Utilizing these relations, we evaluate the truncated differential probability and the maximum average of differential probability of the word-wise Markov (Feistel) ciphers like Rijndael, E2 and the modified version of block cipher E2. This evaluation indicates that all three are provably secure against differential cryptanalysis, and that Rijndael and a modified version of block cipher E2 have stronger security than E2.

keywords. truncated differential cryptanalysis, truncated differential probability, maximum average of differential probability, generalized E2-like transformation, SPN-structure, word-wise Markov cipher, random output-differential cipher

1 Introduction

As a measure of the security of block ciphers, the maximum average of differential probability was defined by Nyberg and Knudsen [15] by generalizing provable security against differential cryptanalysis as introduced by Biham and Shamir [2]. Based on this idea, many new block ciphers have been proposed, e.g. the block cipher MISTY was proposed by M. Matsui [10]. It was designed on the basis of the theory of provable security against differential and linear cryptanalysis.

The block cipher E2 was proposed in [6] as an AES candidate. This cipher uses Feistel structures as a global structure like DES, and uses the SPN (Substitution and Permutation Network)-structure in its S-boxes. [6] said this cipher can be 'proved' to offer immunity against differential cryptanalysis by counting the maximum number of active S-boxes. However, Sugita proposed a method for evaluating the maximux average of differential probability of SPN-structures, and then evaluated the SPN-structure of E2[16, 17]. Using the similar method, Matsui stated that 8-rounds E2 can be defeated by truncated differential cryptanalysis [19, 14], which implies that just counting the maximum number of active S-boxes is not sufficient for proving the security of block ciphers.

The block cipher Rijndael was also proposed as an AES candidate [3]. This cipher uses the SPN (Substitution and Permutation Network)-structure as its basic structure. The basis for proving its security against differential cryptanalysis involves a similar evaluation method as used for E2. Therefore more accurate proof is necessary.

In this paper, we introduce the word-wise Markov (Feistel) cipher and random output-differential (Feistel) cipher as a approximation model for the accurate definition of truncated differential probability, and clarify the relationships among differential, truncated differential and impossible differential cryptanalyses, and propose a new method for evaluating the security of block ciphers against differential, truncated differential and impossible differential cryptanalysis under this model, and propose new structures for block ciphers that are secure against these cryptanalyses. This random output-differential (Feistel) cipher model does not use too strong an assumption because denying this model is equivalent to denying truncated differential cryptanalysis.

This report is organized as follows.

Section 2 defines the structures of word-oriented block ciphers like SPN-Structures, PSN-structures and the E2($'$)-like transformation.

Section 3 defines the differential probability, and defines the word-wise Markov (Feistel) cipher, random output-differential (Feistel) cipher, and using these definitions, defines the truncated differential probability.

Section 4 clarifies the relations between the truncated differential probability and the differential probability of the random output-differential (Feistel) cipher. It then describes a procedure for calculating the truncated differential probability and (maximum average of) the differential probability of typical random output-differential ciphers like SPN-structures including Rijndael and E2($'$)-like transformations. It proves that both Rijndael and the modified E2-like transformation are provably secure against differential, truncated differential and impossible differential cryptanalysis if they can be approximated as random output-differential (Feistel) ciphers.

Section 5 concludes this paper.

2 Structures of Word-oriented Block Ciphers

2.1 Word-oriented Block Ciphers

A word-oriented block cipher is a block cipher whose input and output data is a set of input words of fixed size, and whose operations consist only of word-wise operations of fixed size. In the usual case, the word size is 8, i.e. byte size. Example of these ciphers include Rijndael, E2, etc.

2.2 Feistel Structures

Associate with a function $f : GF(2)^n \to GF(2)^n$, a function $\delta_{2n,f}(L, R) = (R \oplus f(L), L)$ for all $L, R \in GF(2)^n$. $\delta_{2n,f}$ is called the Feistel transformation associated with f. Furthermore, for functions $f_1, f_2, \cdots, f_s : GF(2)^n \to GF(2)^n$, define $\psi_n(f_1, f_2, \cdots, f_s) = \delta_{2n,f_s} \circ \cdots \circ \delta_{2n,f_2} \circ \delta_{2n,f_1}$. We call $D(f_1, f_2, \cdots, f_s) = \psi_n(f_1, f_2, \cdots, f_s)$ as the s-round Feistel structure. At this time, we call the functions $f_1, f_2, \cdots, f_s$ as S-boxes of the Feistel structure $D(f_1, f_2, \cdots, f_s)$.

2.3 SPN-Structures and PSN-Structures

[11] defines SPN-Structures. First we define the 3-layer SPN-structure.

This structure consists of two kinds of layers, i.e. nonlinear layer and bijective linear layer. Each layer has different features as follows.

Nonlinear layer: This layer is composed of m parallel n-bit bijective nonlinear transformations.

Linear layer: This layer is composed of linear transformations over the field $GF(2^n)$ (especially in the case of E2, bit-wise XOR), where inputs are transformed linearly to outputs per word (n-bits).

Next for $s \in \mathbf{N}$ we define the s-layer SPN-structure, which consists of s layers. First is a nonlinear layer, second is a linear layer, third is a nonlinear layer, $\cdots$.

Similarly, for $s \in \mathbf{N}$ we define the s-layer PSN-structure. This layer consists of s layers. First is a linear layer, second is a nonlinear layer, third is a linear layer, $\cdots$.

The SPN-structure is the basic structure of Rijndael, a candidate for AES. We will analyze the security of Rijndael afterwards.

2.4 E2($'$)-like Transformations

[6] proposed the block cipher E2. This cipher has Feistel structures and its S-boxes are composed of 3-layers SPN structures. We generalize this structure, and define E2-like transformations as Feistel structure with S-boxes composed of s-layers (in the case of E2, 3-layers) SPN-structures.

Similarly, we define E2$'$-like transformations as Feistel structures with S-boxes composed of s-layer PSN-structures.

3 Differential Probability, Truncated Differential Probability, Word-wise Markov (Feistel) Cipher and Random Output-Differential (Feistel) Cipher

This section defines the (maximum average of) differential probability, truncated differential probability, word-wise (Feistel) Markov cipher and random output-differential (Feistel) cipher.

3.1 Differential Probability of Block Ciphers

We define the differential of block ciphers. We consider the encryption of a pair of distinct plaintexts by an r-round iterated cipher. Here the round function $Y = f(X, Z)$ is such that, for every round sub-key Z, $f(\ \cdot\ , Z)$ establishes a one-to-one correspondence between the round input X and the round output Y. Let the "difference" ΔX between two plain-texts (or two cipher texts) X and X^* be defined as

$$\Delta X = X \oplus X^*.$$

From the pair of encryption results, one obtains the sequence of differences $\Delta X(0), \Delta X(1), \cdots, \Delta X(r)$ where $X(0) = X$ and $X(0)^* = X^*$ denote the plaintext pair (such that $\Delta X(0) = \Delta X$) and where $X(i)$ and $X^*(i)$ for $(0 < i < r)$ are the outputs of the i-th round, which are also the inputs to the $(i+1)$-th round. The sub-key for the i-th round is denoted as $Z^{(i)}$.

Next we define the i-th round differential and maximum average of differential probabilities.

Definition 1 *[7] An i-round differential is the couple (α, β), where α is the differential of a pair of distinct plaintexts X and X^* and β is a possible difference for the resulting i-th round outputs $X(i)$ and $X^*(i)$. The probability of an i-round differential (α, β) is the conditional probability that β is the difference, $\Delta X(i)$, of the cipher text pair after i rounds given that the plaintext pair (X, X^*) has difference $\Delta X = \alpha$ when the plaintext X and the sub-keys $Z^{(1)}, \cdots, Z^{(i)}$ are independent and uniformly random. We denote this differential probability by $P(\Delta X(i) = \beta | \Delta X = \alpha)$.*

The probability of an s-round differential is known to satisfy the following property.

Lemma 1 *[7] For the Markov cipher, the probability of an s-round differential equals*

$$P(\Delta X(s) = \beta(s) | \Delta X(0) = \beta(0)) = \sum_{\beta(1)} \sum_{\beta(2)} \cdots \sum_{\beta(s-1)} \prod_{i=1}^{s} P(\Delta X(i) = \beta(i) | \Delta X(i-1) = \beta(i-1)).$$

We define the maximum average of differential probability as follows. This value is known to be the best measure with which to confirm that block ciphers are secure against differential cryptanalysis.

Definition 2 *[15] We define the maximum average of differential probability* $ADP^{(s)}_{\max}$ *by*

$$\mathrm{ADP}^{(s)}_{\max} = \max_{\alpha \neq 0, \beta} P(\Delta X(s) = \beta | \Delta X = \alpha).$$

3.2 Word-wise Markov (Feistel) Cipher

[5] uses the truncated differential for the cryptanalysis of word-oriented block ciphers. However, the accurate definition of truncated differential probability is not offered because this cryptanalysis is essentially based on approximation. In this subsection, in order to legitimate this notion, we redefine the truncated differential probability of word-oriented block ciphers.

We consider the encryption of a pair of distinct plaintexts by an r-round iterated cipher. Here the round function $Y = f(X, Z)$ is such that, for every round sub-key $Z = (Z_1, Z_2, \cdots, Z_{m'}) \in GF(2^n)^{m'}$, $f(\,\cdot\,, Z)$ establishes a one-to-one correspondence between the round input $X = (X_1, X_2, \cdots, X_m) \in GF(2^n)^m$ and the round output $Y = (Y_1, Y_2, \cdots, Y_m) \in GF(2^n)^m$.

We define a characteristic function $\chi : GF(2^n)^m \to GF(2)^m$, $(x_1, \cdots, x_m) \longmapsto (y_1, \cdots, y_m)$ by

$$y_i = \begin{cases} 0 & \text{if } x_i = 0 \\ 1 & \text{otherwise,} \end{cases}$$

Hereafter, we call $\chi(x)$ as a characteristic of $x \in GF(2^n)^m$.

For the definition of the truncated differential probability, we define the word-wise Markov cipher as a real block-cipher model, in the same way as the Markov cipher was in [7]

Definition 3 *A word-oriented cipher with round function* $Y = f(X, Z) (X = (X_1, X_2, \cdots, X_m) \in GF(2^n)^m,\ Y = (Y_1, Y_2, \cdots, Y_m) \in GF(2^n)^m, Z = (Z_1, Z_2, \cdots, Z_{m'}) \in GF(2^n)^{m'})$*, is a word-wise Markov cipher if for all choices of* $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_m) \in GF(2^n)^m (\alpha \neq 0)$*,* $\beta = (\beta_1, \beta_2, \cdots, \beta_m) \in GF(2^n)^m (\beta \neq 0)$ *and* $p \in \{1, 2, \cdots, m'\}$*,*

$$P(\Delta Y_p = \beta_p | \Delta X = \alpha, X = \gamma)$$

is independent of γ*, and* $P(\Delta Y_p = \beta_p | \Delta Y_p \neq 0, \Delta X = \alpha, X = \gamma)$ $(p = 1, 2, \cdots, m)$ *are jointly statistically independent when the sub-key* Z *is uniformly random, or, equivalently, if*

$$P(\Delta Y_p = \beta_p | \Delta X = \alpha, X = \gamma) = P(\Delta Y_p = \beta_p | \Delta X = \alpha)$$

for all choices of γ *and* $P(\Delta Y_p = \beta_p | \Delta Y_p \neq 0, \Delta X = \alpha)$ $(p = 1, 2, \cdots, m)$ *are jointly statistically independent when the sub-key* Z *is uniformly random, where* $\Delta X = (\Delta X_1, \Delta X_2, \cdots, \Delta X_m)$*,* $\Delta Y = (\Delta Y_1, \Delta Y_2, \cdots, \Delta Y_m)$ *are the differential of* X*,* Y*, respectively.*

Example. the PSN-structure is a word-wise Markov cipher, if every bijective nonlinear function in a nonlinear layer consists of a concatenation of XOR and substitution (like DES does). Therefore, block cipher Rijndael and the S-boxes of block cipher E2 are also word-wise Markov ciphers with the same kind of nonlinear functions.

We expand this definition to the Feistel cipher.

Definition 4 *We define a word-wise Markov Feistel cipher as a Feistel cipher whose S-boxes are word-wise Markov ciphers.*

Example. E2′-like transformation is a word-wise Markov Feistel cipher because the PSN-structure is a word-wise Markov cipher if every nonlinear function in a nonlinear layer consists of the concatenation of XOR of the key and substitution (like DES does).

3.3 Random Output-Differential (Feistel) Cipher

As preparation for defining the random output-differential cipher, we define the random output-differential transformation.

Definition 5 *A word-oriented transformation* $Y = g(X, Z)$ $(X = (X_1, X_2, \cdots, X_m) \in GF(2^n)^m$, $Y = (Y_1, Y_2, \cdots, Y_m) \in GF(2^n)^m$, $Z = (Z_1, Z_2, \cdots, Z_{m'}) \in GF(2^n)^{m'})$, *is a random output-differential transformation, if for any input-differential value* α, *the following relation is satisfied,*

$$P(\Delta Y = \beta | \Delta X = \alpha) = p^{\mathrm{h}(\chi(\beta))} P(\chi(\Delta Y) = \chi(\beta) | \Delta X = \alpha),$$

when keys are randomly selected, where h *is the function that indicates the Hamming weight of the input value,* $p = 1/(2^n - 1)$, *and* $\Delta X = (\Delta X1, \Delta X_2, \cdots, \Delta X_m)$, $\Delta Y = (\Delta Y_1, \Delta Y_2, \cdots, \Delta Y_m)$ *are the differential of* X, Y, *respectively.*

Using this definition, we define the random output-differential cipher for word-oriented block cipher as approximation model of word-wise Markov cipher.

Definition 6 *A word oriented cipher with round functions* $X(i+1) = f(X(i), Z^{(i)})(i = 0, 1, \cdots, r-1)$, *where* $Z^{(i)}(i = 0, 1, \cdots)$ *are sub-keys, is a random output-differential cipher if for any random output-differential transformation* $X(0) = g(X, Z^{(0)})$, *the composite transformation* $X(1) = f(g(X, Z^{(0)}), Z^{(1)})$ *is also a random output-differential transformation.*

At this time, we call a round function which composes a random output-differential cipher by concatenating, as random output-differential round function.

We expand this definition to the Feistel cipher.

Definition 7 *A Feistel cipher with S-boxes* $Y = f(X, Z^{(i)})$ $(i = 0, 1, \cdots)$, *where* $Z^{(i)}(i = 0, 1, \cdots)$ *are sub-keys and i-th round output is* $X(i) = (X(i)_L, X(i)_R)$, *is a random output-differential Feistel cipher, if its S-boxes are random output-differential ciphers and the round function of the Feistel cipher*

$$(X(i+1)_L, X(i+1)_R) = (X(i)_R, X(i)_L \oplus f(X(i)_R, Z^{(i)}))$$

is a random output-differential round function.

Matsui stated in his presentation of [14] that 8-round E2 can be cryptanalyzed by truncated differential cryptanalysis only assuming randomness of keys. However, this is not accurate, because he tacitly assumes this random output-differential cipher as an approximation model of E2 in his explanation.

However, this approximation may be effective for word-wise Markov (Feistel) cipher like E2, E2′-like transformation and Rijndael. In fact, in the case of E2′-like transformation with 2-layer PSN-structures, which is also a word-wise Markov Feistel cipher for example, let the $\Delta X \in GF(2^n)^{2m}$ be a input differential of this cipher, if the input-differential of S-box $\Delta W = (\Delta W_1, \Delta W_2, \cdots, \Delta W_m) \in GF(2^n)^m (\chi(\Delta W) = \gamma' \in GF(2)^m)$ is randomly distributed with the probability $P(\Delta W = \gamma | \chi(\Delta W) = \gamma', \Delta X = \alpha) = p^{\mathrm{h}(\gamma')}$ for all $\gamma = (\gamma_1, \gamma_2, \cdots, \gamma_m)$ (where $\chi(\gamma) = \gamma'$), then the output-differential of S-box $\Delta U = (\Delta U_1, \Delta U_2, \cdots, \Delta U_m) \in GF(2^n)^m$ $(\chi(\Delta U) = \delta' \in GF(2)^m)$ is supposed to be approximately random, i.e. approximately $P(\Delta U = \delta | \chi(\Delta U) = \delta', \Delta X = \alpha) = p^{\mathrm{h}(\delta')}$, where $\delta = (\delta_1, \delta_2, \cdots, \delta_m), \chi(\delta) = \delta'$ because, for the input-differential of nonlinear layer $\Delta W = (\Delta W_1, \Delta W_2, \cdots, \Delta W_m) \in GF(2^n)^m$, each $P(\Delta W_p = \gamma_p | \chi(\Delta W) = \gamma', \Delta X = \alpha) = p = 1/(2^n - 1)$ implies $P(\Delta U_p = \delta_p | \chi(\Delta U) = \delta', \Delta X = \alpha) = p = 1/(2^n - 1)$ and each $P(\Delta U_p = \delta_p | \Delta W_p = \gamma_p \neq 0, \Delta X = \alpha)$ $(\gamma = (\gamma_1, \gamma_2, \cdots, \gamma_m), \chi(\gamma) = \gamma')$ are jointly statistically independent from the definition of word-wise Markov cipher.

So we use this random output-differential cipher as an effective approximation model in the following discussion.

Note. Matsui assumed the randomness of input-differential of nonlinear layers $\Delta W = (\Delta W_1, \cdots, \Delta W_m)$, i.e.

$$P(\Delta W = \gamma | \chi(\Delta W) = \gamma', \Delta X = \alpha) = p^{\mathrm{h}(\gamma')} P(\chi(\Delta W) = \gamma' | \Delta X = \alpha)$$

instead of the randomness of output-differential of nonlinear layers $\Delta U = (\Delta U_1, \cdots, \Delta U_m)$, i.e.

$$P(\Delta U = \beta | \chi(\Delta U) = \beta', \Delta X = \alpha) = p^{\mathrm{h}(\beta')} P(\chi(\Delta U) = \beta' | \Delta X = \alpha)$$

in his presentation of [14]. The randomness of ΔW is a stronger assumption than the randomness of ΔU, because, in the case of E2′-like transformation with 2-layer PSN-structures for example, the randomness of ΔW also implies the randomness of ΔU. Furthermore, the randomness of ΔW may be too strong or even nonsense, because the randomness of input-differentials of linear layer $\Delta W = (\Delta W_1, \cdots, \Delta W_m)$ do not always yield the randomness of input-differentials of nonlinear layer $\Delta U = (\Delta U_1, \cdots, \Delta U_m)$: Two input-differential words of nonlinear layers ΔW_{p_1}, ΔW_{p_2} $(p_1 \neq p_2)$ may be both random, i.e.

$$P(\Delta W_{p_1} = \gamma_{p_1} | \Delta X = \alpha) = P(\Delta W_{p_2} = \gamma_{p_2} | \Delta X = \alpha) = p = 1/(2^n - 1)$$

but coincide, i.e. constantly $\Delta W_{p_1} = \Delta W_{p_2}$.

Therefore, we interpret Matsui's tacit assumption in his explanation as a random output-differential (Feistel) cipher.

3.4 Truncated Differential Probability

Using these definitions, we can accurately define the truncated differential probability. In this definition, as a cipher model, we consider a cipher with a random output-differential initial transformation $X(0) = g(X, Z^{(0)})$, and a random output-differential round function $X(i+1) = f(X(i), Z^{(i)})(i = 0, 1, \cdots, r-1)$ where $Z^{(i)}(i = 0, 1, \cdots)$ are sub-keys.

Definition 8 *Let $X(0) = g(X, Z^{(0)})$ be an arbitrary random output-differential initial transformation and $X(i+1) = f(X(i), Z^{(i)})$ be a round function such that $X(r) = (f \circ \cdots \circ f \circ g)(X, Z^{(0)}, Z^{(1)}, \cdots, Z^{(r)})$ is also a random output-differential cipher for all r. An i-round truncated differential of i-round iterated cipher $X(r) = (f \circ \cdots \circ f)(X(0), Z^{(1)}, \cdots, Z^{(r)})$ is the couple (α', β'), where α is the differential of a pair of distinct values $X(0)$ and $X^*(0)$, $\alpha' = \chi(\alpha)$ is the characteristic of α; β is a possible difference for the resulting i-th round outputs $X(i)$ and $X^*(i)$; $\beta' = \chi(\beta)$ is the characteristic of β. The probability of i-round truncated differential (α', β') is the conditional probability that β' is the characteristic of difference $\Delta X(i)$ of the cipher text pair after i rounds given that the characteristic of pair $(X(0), X(0)^*)$ has difference $\chi(\Delta X) = \alpha'$ when the plaintext X and the sub-keys $Z^{(0)}, \cdots, Z^{(i)}$ are independent and uniformly random. We denote this truncated differential probability by $P'_i(\beta'(i), \beta'(0)) = P(\chi(\Delta X(i)) = \beta'(i) | \chi(\Delta X(0)) = \beta'(0), \Delta X = \alpha)$.*

This definition is well defined if we assume the random output-differential (Feistel) cipher. Without the assumption, this is not well-defined, because two input-differential values with same characteristic value do not always yield the same truncated differential probabilities. We assume this model as an effective approximation model of a word-wise Markov cipher.

4 Truncated Differential Probability and Differential Probability of Random Output-Differential (Feistel) Ciphers

4.1 Truncated Differential Probability of PSN-structures and Differential Probability of SPN-structures

In this subsection, we evaluate the truncated differential probability of the $2s$ layer PSN-structure and (the maximum average of) the differential probability of the $(2s+1)$ layer SPN-structure, where we assume all random functions are bijective. In this calculation, we first calculate the truncated differential probability of the $2s$ layer PSN-structure, and, using this probability, we calculate (the maximum average of) the differential probability of the $(2s+1)$ layer SPN-structure.

We assume the first nonlinear layer is a random output-differential (initial) transformation, and the round functions, which are composed of a linear layer and a nonlinear layer, i.e. 2-layer PSN-structures, is a random output-differential round function. We denote ΔX as the input-differential of the first nonlinear layer, $\Delta X(0)$ as the output-differential of the first nonlinear layer, $\Delta X(1)$ as the output-differential of the second nonlinear layer, $\cdots$, $\Delta X(s)$ as the output-differential of $(s+1)$-th nonlinear layer, $\Delta Y(0)$ as the input-differential of the first nonlinear layer, $\Delta Y(1)$ as the input-differential of the second nonlinear layer, $\cdots$, $\Delta Y(s)$ as the input-differential of the $(s+1)$-th nonlinear layer.

We denote the differential probability of $(2s+1)$-layer SPN-structures as

$$P_i(\beta(i), \alpha) = P(\Delta X(i) = \beta(i) | \Delta X = \alpha).$$

We denote the truncated differential probability of $2s$-layer PSN-structures as

$$P'_i(\beta'(i), \beta'(0)) = P(\chi(\Delta X(i)) = \beta'(i) | \chi(\Delta X(0)) = \beta'(0), \Delta X = \alpha).$$

The relation between differential probability and truncated differential probability can be represented as follows, where $\beta'(i) = \chi(\beta(i))$ for all $i = 0, 1, \cdots, s$,

$$\begin{aligned} &P_i(\beta(i), \alpha) = \\ &\quad \sum_{\beta'(0)} P(\Delta X(i) = \beta(i)) | \chi(\Delta X(i)) = \beta'(i), \Delta X = \alpha) \\ &\quad * \ \ P'_i(\beta'(i), \beta'(0)) * P(\chi(\Delta X(0) = \beta'(0)) | \Delta X = \alpha). \end{aligned}$$

In this case, if we assume the initial transformation is a random output-differential transformation and

$$P(\chi(\Delta X(0)) = \beta'(0) | \Delta X = \alpha) = \begin{cases} 1 & \text{if } \beta'(0) = \chi(\alpha) \\ 0 & \text{otherwise,} \end{cases}$$

as a natural approximation model of the first nonlinear layer, we can prove

$$P_i(\beta(i), \alpha) = p^{\mathrm{h}(\beta'(i))} P'_i(\beta'(i), \alpha'),$$

because

$$P(\Delta X(i) = \beta(i) | \chi(\Delta X(i)) = \beta'(i), \Delta X = \alpha) = p^{\mathrm{h}(\beta'(i))},$$

from the assumption of random output-differential cipher, where $p = 1/(2^n - 1)$ and h is the function that indicates the Hamming weight of the input value.

This relation clearly indicates the relationship between the differential probability and the truncated differential probability. From this relation we can easily calculate the differential probability from the truncated differential probability in the case of random output-differential cipher. This relation also implies that the possibility of truncated differential cryptanalysis is equivalent to the possibility of differential cryptanalysis, because the ratio of obtained probability to average probability do not change.

4.2 Procedure for Calculating Differential and Truncated Differential Probability of the SPN-structure

The procedure for calculating truncated differential probability and the maximum average of the differential probability in case of the SPN structure is as follows.

For this procedure, we define function $\mathrm{N}(P, \gamma, \delta)$ for $m \times m$ matrix P over $GF(2^n)$ and $\gamma, \delta \in GF(2)^m$ by

$$\begin{aligned} \mathrm{N}(P, \gamma, \delta) = \#\{&(\Delta X, \Delta Y) \in (GF(2^n)^m)^2 \setminus \{0\} | \\ &\Delta Y = P\Delta X, \chi(\Delta X) = \gamma, \chi(\Delta Y) = \delta\}, \end{aligned}$$

For this calculation we define semi-order $\prec$ in $GF(2)^m$ as follows.

$$a \prec b \Leftrightarrow (\forall i; (a(i) = 1 \Rightarrow b(i) = 1)) \wedge (a \neq b)$$

where we denote $a(i)$ and $b(i)$ as the i-th significant bits of a and b, respectively.

For $m \times m$ matrix P over $GF(2^n)$ and $\gamma, \delta \in GF(2)^m$, we define

$$\begin{aligned}\mathrm{M}(P, \gamma, \delta) = \#\{(\Delta X, \Delta Y) \in (GF(2^n)^m)^2 \setminus \{0\} | \\ \Delta Y = P\Delta X, \chi(\Delta X) \preceq \gamma, \chi(\Delta Y) \preceq \delta\},\end{aligned}$$

and $\mathrm{N}(P, \gamma, \delta)$ can be calculated recursively, using the following relations.

$$\mathrm{N}(P, \gamma, \delta) = \mathrm{M}(P, \gamma, \delta) - \sum_{(\gamma', \delta') \prec (\gamma, \delta)} \mathrm{N}(P, \gamma', \delta')$$

In this case, we assume a random output-differential cipher. Under this assumption, we can prove the following lemma.

Lemma 2

$$\begin{aligned}&P'_i(\beta'(i), \beta'(0)) = \\ &\sum_{\beta'(i-1)} \mathrm{N}(P, \beta'(i), \beta'(i-1)) p^{\mathrm{h}(\beta'(i-1))} P'_{i-1}(\beta'(i-1), \beta'(0)),\end{aligned}$$

where $p = 1/(2^n - 1)$.

Proof. From the assumption of a random output-differential cipher,

$$\begin{aligned}&P(\Delta X(i-1) = \beta(i-1) | \chi(\Delta X(0)) = \beta'(0), \Delta X = \alpha) \\ &= P(\Delta X(i-1) = \beta(i-1) | \chi(\Delta X(i-1)) = \beta'(i-1), \Delta X = \alpha) \\ &* \quad P(\chi(\Delta X(i-1)) = \beta'(i-1) | \chi(\Delta X(0)) = \beta'(0), \Delta X = \alpha) \\ &= p^{\mathrm{h}(\beta'(i))} P(\chi(\Delta X(i-1)) = \beta'(i-1) | \chi(\Delta X(0)) = \beta'(0), \Delta X = \alpha) \\ &= p^{\mathrm{h}(\beta'(i-1))} P'_{i-1}(\beta'(i-1), \beta'(0)),\end{aligned}$$

where $\beta'(i-1) = \chi(\beta(i-1))$.

From the definition of N,

$$\begin{aligned}\mathrm{N}(P, \beta'(i), \beta'(i-1)) = \#\{(\Delta X(i), \Delta X(i-1)) \in (GF(2^n)^m \setminus \{0\})^2 | \\ \Delta X(i) = P\Delta X(i-1), \chi(\Delta X(i)) = \beta'(i), \chi(\Delta X(i-1)) = \beta'(i-1)\},\end{aligned}$$

.

Therefore,

$$\begin{aligned}&P'_i(\beta'(i), \beta'(0)) \\ &= \sum_{\beta'(i-1)} \mathrm{N}(P, \beta'(i), \beta'(i-1)) P(\Delta X(i-1) = \beta(i-1) | \chi(\Delta X(0)) = \beta'(0)) \\ &= \sum_{\beta'(i-1)} \mathrm{N}(P, \beta'(i), \beta'(i-1)) p^{\mathrm{h}(\beta'(i-1))} P'_{i-1}(\beta'(i-1), \beta'(0)),\end{aligned}$$

This lemma, yields the following procedure.

1) **Computing the Number of All Differential Paths**
   For given P, calculate $N(P, \gamma, \delta)$ for every $\gamma, \delta \in GF(2)^m$.
   $\mathrm{M}(P, \gamma, \delta)$ can be easily calculated by simple rank calculation as follows.

$$
\begin{aligned}
&\mathrm{M}(P,\gamma,\delta)\\
&= \#\{(\Delta X,\Delta Y)\in GF(2^n)^{2m}\setminus\{0\} | P\Delta X=\Delta Y, F(\bar{\gamma})\Delta X=0, F(\bar{\delta})\Delta Y=0\}\\
&= 2^{n*\dim\{(\Delta X,\Delta Y)\in GF(2)^{2m}\setminus\{0\}|P\Delta X=\Delta Y,F(\bar{\gamma})\Delta X=0,F(\bar{\delta})\Delta Y=0\}}-1\\
&= 2^{n(2m-\mathrm{rank}\left(\begin{array}{cc} P & E\\ F(\bar{\gamma}) & O\\ O & F(\bar{\delta})\end{array}\right))}-1,
\end{aligned}
$$

where $\bar{\gamma}$ and $\bar{\delta}$ are the complements of γ and δ, respectively, E is an identity matrix, and $F(\bar{\gamma})$, $F(\bar{\delta})$, denote the diagonal matrices over $GF(2^n)$ whose (i,i) component equals the i-th significant bit of $\bar{\gamma}$, $\bar{\delta}$ for $i=1,\cdots,m$, respectiveley.

$\mathrm{N}(P,\gamma,\delta)$ can be calculated recursively from the values of $\mathrm{M}(P,\gamma,\delta)$, using the following relation.

$$
\mathrm{N}(P,\gamma,\delta)=\mathrm{M}(P,\gamma,\delta)-\sum_{(\gamma',\delta')\prec(\gamma,\delta)}\mathrm{N}(P,\gamma',\delta')
$$

2) **Initialization**
For given $\alpha'\in GF(2)^m$, calculate $P'_0(\beta'(0),\alpha')$ for every $\beta'(0)\in GF(2)^m$, where

$$
P'_0(\beta'(0),\alpha')=\begin{cases}1 & \text{if } \beta'(0)=\alpha'\\ 0 & \text{otherwise,}\end{cases}
$$

3) **Recursive Computation of Truncated Differential Probability**
Utilizing the values of $N(P,\gamma,\delta)$, calculate $P'_i(\beta'(i),\alpha')$ recursively for every $\beta'(i)\in GF(2)^m$.

$$
\begin{aligned}
&P'_i(\beta'(i),\alpha')=\\
&\sum_{\beta'(i-1)}\mathrm{N}(P,\beta'(i),\beta'(i-1))p^{\mathrm{h}(\beta'(i-1))}P'_{i-1}(\beta'(i-1),\alpha')
\end{aligned}
$$

4) **Calculation of (Maximum Average of) Differential Probability**
Evaluate $P_i(\beta(i),\alpha)$ by

$$
P_i(\beta(i),\alpha)=p^{\mathrm{h}(\beta'(i))}P'_i(\beta'(i),\alpha')
$$

With this procedure we can compute the truncated differential probability of PSN-structures and (the maximum average of) the differential probability of SPN-structures with 16 input words. Furthermore, applying this procedure to the each MixColumn transformations of Rijndael, allows us to compute the truncated differential probability and (the maximum average of) the differential probability. From this computation, the maximum average of the differential probability of 7-layer Rijndael including 4 nonlinear layers, i.e. 4-round Rijndael, is upper-bounded by $1.00*p^{16}$ $(=1.065*2^{-128})$ and that of 9-layer Rijndael including 5 nonlinear layers, i.e. 5-round Rijndael, is upper-bounded by $0.940*p^{16}$ $(=1.0007*2^{-128})$ [1]. To be secure against differential and truncated differential cryptanalysis, 2 more layers (1 round) are necessary to avoid the exhaustive search of the the last 2 layers (1 round). This implies a total of 80 S-boxes is needed.

[1] [12] stated that 5-round differential with probability $1.06*2^{-128}$ was found, but this was typo. The correct round is 4.

4.3 Truncated Differential Probability of E2′-like Transformation

Using the values of the differential probability of the $2r$-layer PSN-structures, we can calculate the truncated differential probability of E2′-like transformations recursively. In this calculation, we assume the random output-differential Feistel cipher, hence the probabilities for $\chi(\Delta x \oplus \Delta y) = 1$ and $\chi(\Delta x \oplus \Delta y) = 0$ for two random output-differential values $\Delta x, \Delta y, \in GF(2^n) \setminus \{0\}$ are $(2^n - 2)/(2^n - 1)$ and $1/(2^n - 1)$, respectively.

The procedure for calculating the truncated differential probability of the E2′-like transformation is as follows.

1) **Computation of Truncated Differential Probability of Round Functions**
Using the procedure for calculating truncated differential probability of $2r$-layer PSN-structure, calculate the truncated differential of round functions. Hereafter, we denote the truncated differential probability of the i-th round function for the truncated differential $(\zeta'(i), \zeta'(i-1))$ by $Q'_r(\zeta'(i), \zeta'(i-1)) = P(\chi(\Delta X(i)) = \zeta'(i) | \chi(\Delta X(i-1)) = \zeta'(i-1))$ for $\zeta'(i), \zeta'(i-1) \in GF(2)^m$

2) **Initialization**
Let $\zeta'(0) = (\Delta L'(0), \Delta R'(0)) \in GF(2)^{2m}$. For given $\eta' \in GF(2)^{2m}$, calculate $P'_0(\zeta'(0), \eta')$ for every $\zeta'(0) \in GF(2)^{2m}$, where we assume

$$P'_0(\zeta'(0), \eta') = \begin{cases} 1 & \text{if } \zeta'(0) = \eta' \\ 0 & \text{otherwise,} \end{cases}$$

3) **Recursive Computation of Truncated Differential Probability**
Let $\zeta'(i) = (\Delta L'(i), \Delta R'(i)) \in GF(2)^{2m}$, $\zeta(i) = (\Delta L(i), \Delta R(i)) \in GF(2)^{2m}$, where $\chi(\zeta(i)) = \zeta'(i), \chi(\Delta L(i)) = \Delta L'(i), \chi(\Delta R(i)) = \Delta R'(i)$. Utilizing the values of truncated differential probabilities of round functions, calculate $P'_i(\zeta'(i), \eta')$ recursively for every $\zeta'(i) \in GF(2)^{2m}$.

$$P'_i(\zeta'(i), \eta') = \sum_{\substack{\xi', \\ \chi(L(i-1) \oplus \xi) = \Delta R'(i), \\ \chi(\xi) = \xi'}} Q'_i(\xi', \Delta R'(i-1)) P'_{i-1}(\zeta'(i-1), \eta')$$

4) **Calculation of (Maximum Average of) Differential Probability**
Calculate $P_i(\zeta(i), \eta)$ by

$$P_i(\zeta(i), \eta) = p^{\mathrm{h}(\zeta'(i))} P'_i(\zeta'(i), \eta'),$$

where $\chi(\eta) = \eta'$.

4.4 (Maximum Average of) Differential and Truncated Differential Probability of E2′-like Transformation

In this subsection, we evaluate the maximum average of the differential probability of E2′-like transformations with proper initial transformations, where we assume the all linear layers are same as that of E2.

First we consider E2′-like transformations with 2-layer PSN-structures. In this case, a nonlinear layer with 16 nonlinear functions, or 2-round E2′-like transformations with 2-layer PSN-structures can be adopted as the approximately random output-differential initial transformation. 8-round E2′-like transformation with 2-layer PSN-structures with proper initial transformation has maximum average of differential probability of less than $0.940 * p^{16}$ $(= 1.0007 * 2^{-128})$. In this case, it is provably secure with 80 nonlinear functions. To offer security against differential cryptanalysis, 2 more rounds are necessary, which means it needs a total of 96 nonlinear functions.

If we slightly change linear transformation of SPN-structures, it can be provably secure with 72 nonlinear functions. To offer security against differential cryptanalysis, 2 more rounds are necessary, which means it needs a total of 88 nonlinear functions.

Next we consider E2′-like transformations with 4-layer PSN-structures. In this case, a nonlinear layer with 16 nonlinear functions or 2-round E2′-like transformations with 2-layer or 4-layer PSN-structures can be adopted as the proper initial transformation. A 5-round E2′-like transformation with 4-layer PSN-structures with proper initial transformation has maximum average of differential probability lower than $0.940 * p^{16}$ $(= 1.0007 * 2^{-128})$. In this case, it is provably secure with 96 nonlinear functions. To be secure against differential cryptanalysis, 1 more round is necessary, which means it needs a total of 112 nonlinear functions to avoid the exhaustive search of the final round.

On the other hand, an 8-round E2-like transformation with 3-layer SPN-structures, has maximum average of differential probability lower than $0.940 * p^{16}$ $(= 1.0007 * 2^{-128})$. In this case, it is provably secure with 128 S-boxes (in this case, approximately random output-differential initial function is not necessary because of the first nonlinear layers of the first and second S-boxes). To be secure against differential cryptanalysis, 1 more round is necessary, considering the exhaustive search of the final round, which implies it needs a total of 144 S-boxes.

These results means that E2′-like transformations with 2-layer PSN-structures is more secure than 3 or 4 layer.

The block cipher MISTY with 16-input words and 3-rounds has maximum average of differential probability equal to $p_{\max}^{16}$, where $p_{\max}$ is the maximum average of differential probability of nonlinear functions. In this case, it is provably secure with 81 S-boxes. To be secure against differential cryptanalysis, 1 more round is necessary, which implies it needs a total of 108 S-boxes.

4.5 Impossible Differential Cryptanalysis of Rijndael, E2′-like Transformation

Impossible differential cryptanalysis is a cryptanalysis against block ciphers which utilizes the pair of input and output-differentials whose differential probability equals 0 [1].

In the previous procedure, we proposed the procedure which calculates the truncated differential probability of random output-differential (Feistel) ciphers. It follows that from the relations between truncated differential probability and differential probability we can also calculate the differential probability.

From the values of the differential probability, our procedure can calculate the resistance against impossible differential cryptanalysis, by counting the number of differentials whose probabilities equal 0. In the case of E2′-like transformations with 2-layer PSN-structures, it can be proved that 9-rounds offer security against impossible differential cryptanalysis while 8-rounds do not. In the case of E2-like transformations with 3-layer SPN-structures, 8-rounds offer security against impossible differential cryptanalysis and 7-rounds do not. Comparing the numbers of nonlinear functions, E2′-like transformations with 2-layer PSN-structures is superior to E2-like transformations with 3-layer SPN-structures, i.e. the basic structure of block cipher E2.

In the case of Rijndael, it can be proved that 7-layers (including 4 nonlinear layers) offers security against impossible differential cryptanalysis while 5-layers (including 3 nonlinear layers) do not. Comparing the numbers of nonlinear functions, basic structure of Rijndael has a little higher level of security against impossible cryptanalysis than E2′-like transformation with 2-layer PSN-structures. However, considering the amount of linear layer operations, E2′-like transformations with 2-layer PSN-structures may be superior to the basic structure of Rijndael, because the linear layer of E2′-like transformations consists of only "xor" whereas that of Rijndael consists of heavier linear transformation over Galois field $\mathrm{GF}(2^8)$.

5 Conclusion

This paper examined the truncated differential probability and the differential probability of the word-oriented Markov ciphers and random output-differential (Feistel) ciphers like Rijndael and (modified) E2 and clarified the relations among the differential, truncated differential and the impossible differential cryptanalysis of the random output-differential (Feistel) cipher. This random output-differential (Feistel) cipher uses a weaker assumption than the assumption that all S-box differentials are equally likely. This is not a strong assumption because denying this model is equivalent to denying the truncated differential cryptanalysis. We then described a procedure for calculating the truncated differential probability and (maximum average of) the differential probability of such ciphers. Using this procedure, we computed and proved the security of Rijndael, E2 and the E2′-like transformation against differential, truncated differential and impossible differential cryptanalyses under the assumption of a random output-differential (Feistel) cipher. Our evaluation finds that Rijndael is the most secure, and the E2′-like transformation with 2-layer PSN structure is a little less secure. However, the linear transformation in E2′-like transformations is lighter than that of Rijndael and can be improved by slightly changing, so the overall speed may be the highest (may be "not" the highest). Our results implies that SPN-structures (like Rijndael, Serpent) and Feistel structures with S-boxes composed of 2-layer PSN-structures (like E2-like transformation with 2-layer PSN-structures) have no disadvantage in terms of security against differential and truncated differential cryptanalysis. We can similary evaluate the security of Feistel structures with S-boxes composed of 2-layer SPN-structures (like Twofish [18]) against differential and truncated differential cryptanalysis, though we have not evaluated Twofish yet because Twofish is not composed of just word-wise operations of fixed size. However, Feistel structures with 2-layer SPN-structures can be proved to be secure and have no disadvantage in terms of security against differential and truncated differential cryptanalysis, if we select the proper linear transformations in their SPN-structures.

References

[1] E.Biham, A.Biryukov and A.Shamir, "Cryptanalysis of Skipjack Reduced to 32 Rounds Using Impossible Differentials." J. Stern (Ed.): EUROCRYPT'99, LNCS 1592, pp. 12-23, Springer-Verlag, Berlin, 1999.

[2] E.Biham and A.Shamir. "Differential Cryptanalysis of DES-like Cryptosystems." Journal of Cryptology, Vol.4, No.1, pp.3-72, 1991. (The extended abstract was presented at CRYPTO'90).

[3] J. Daemen and V. Rijmen. "AES Proposal Rijndael," AES Round 1 Technical Evaluation CD-1: Documentation, National Institute of Standards and Technology, Aug 1998. See **http://www.nist.gov/aes**.

[4] T.Jakobsen and L.R.Knudsen. "The Interpolation Attack on Block Cipher." In E.Biham, editor, Fast Software Encryption - 4th International Workshop, FSE'97, Volume 1267 of Lecture Notes in Computer Science, pp.28-40, Berlin, Heidelberg, NewYork, Springer-Verlag, 1997.

[5] L.R. Knudsen and T.A. Berson, "Truncated Differentials of SAFER." In Fast Software Encryption - Third International Workshop, FSE'96, Volume 1039 of Lecture Notes in Computer Science, Berlin, Heidelberg, NewYork, Springer-Verlag, 1996.

[6] M. Kanda et al. "A New 128-bit Block Cipher E2" Technical Report of IEICE. ISEC98-12.

[7] X. Lai, J. L. Massey and S. Murphy, "Markov Ciphers and Differential Cryptanalysis," Advances in Cryptography-EUROCRYPTO '91. Lecture Notes in Computer Science, Vol. 576. Springer-Verlag, Berlin, 1992, pages. 86-100.

[8] M.Matusi, "Linear Cryptanalysis Method for DES Cipher." In T. Helleseth, editor, Advances in Cryptology - EUROCRYPT'93, Volume765 of Lecture Notes in Computer

Science, pp.386-397. Springer-Verlag, Berlin, Heidelberg, NewYork, 1994. (A preliminary version written in Japanese was presented at SCIS93-3C).

[9] M. Matsui, "New structure of block ciphers with provable security against differential and linear cryptanalysis," In Dieter Grollman, editor, Fast Software Encryption: Third International Workshop, volume 1039 of Lecture Notes in Computer Science, pages 205-218, Cambridge, UK, 21-23 February 1996. Springer-Verlag.

[10] M. Matsui, "New block encryption algorithm MISTY." In Eli Biham, editor, Fast Software Encryption: 4th International Workshop, volume 1267 of Lecture Notes in Computer Science, pages 54-68, Haifa, Israel, 20-22 January 1997. Springer-Verlag

[11] A. J. Menezes, P. C. van Oorschot, S. A. Vanstone, "Handbook of Applied Cryptography", CRC Press, 250-250 (1997).

[12] Moriai, S., Sugita, M., Aoki, K., Kanda, M. "Security of E2 against truncated Differential Cryptanalysis" Sixth Annual Workshop on Selected Areas in Cryptography (SAC'99), pages, 133-143 (1999).

[13] Moriai, S. et al. "Security of E2 against truncated Differential Cryptanalysis" Technical Report of IEICE, to appear.

[14] Matsui, M. and Tokita, T. "Cryptanalysis of a Reduced Version of the Block Cipher E2" in 6-th international workshop, preproceedings FSE'99

[15] K. Nyberg and L. R. Knudsen, "Provable security against a differential attack," in Advances in Cryptology - EUROCRYTO'93, LNCS 765, pages 55-64, Springer-Verlag, Berlin, 1994.

[16] M. Sugita, "Security of Block Ciphers with SPN-Structures" Technical Report of IEICE. ISEC98-30.

[17] M. Sugita, K. Kobara, H. Imai, "Pseudorandomness and Maximum Average of Differential Probability of Block Ciphers with SPN-Structures like E2." Second AES Workshop, 1999.

[18] B. Schneier, J. Kelsey, D. Whiting, D. Wagner and C. Hall, "Twofish: A 128-Bit Block Cipher," AES Round 1 Technical Evaluation CD-1: Documentation, National Institute of Standards and Technology, June 1998. See **http://www.nist.gov/aes**.

[19] T. Tokita, M. Matsui, "On cryptanalysis of a byte-oriented cipher", The 1999 Symposium on Cryptography and Information Security, pages 93-98 (In Japanese), Kobe, Japan, January 1999.